1. a. Approximately how many floating point operations (flops) will be required to solve a $25,000 \times 25,000$ linear system by Gaussian elimination?

**solution:**

$\mathbf{A}$ is $25,000,000$ ($n = 25000$). Gaussian elimination "costs" about $(2/3)n^3$ flops:

$$\text{cost} \doteq \frac{2}{3}(25000)^3 \doteq 10.4 \times 10^{12} \;\; \text{flops} \;\; \doteq 10 \;\; \text{teraflops}$$

b. Assuming one used a computer capable of executing a sustained 80 million floating point operations per second (80Mflops), approximately how long would it take to solve this system?

**solution:**

Computer executes about 80 Mflops $= 80 \times 10^6$ flops/sec.

Time Required:

$$\frac{10.4 \times 10^{12}\text{flops}}{80 \times 10^6\text{flops/sec}} \doteq 130,000 \;\; \text{sec} \;\; \doteq 1\,\frac{1}{2} \;\; \text{days}$$

c. What would be the minimum available RAM required on this system in order not to have to page out to swap during execution of this solution?

**solution:**

$\mathbf{A}$ is $n \times n$, has $n^2$ elements:

$$(25000)^2 = 625 \times 10^6 \text{ elements}$$

Double precision (i.e. MATLAB precision) requires 8 bytes per element:

$$8 * (625 \times 10^6 \text{ elements}) = 5 \times 10^9 \text{ elements} \doteq 5 \text{ GB}$$

2. A particular PC, using Gaussian elimination, solves a system of $1,000$ equations in $1,000$ unknowns in about two seconds, with no apparent signs of excessive paging or swapping. Approximately how long would it this same PC to solve a $3,000 \times 3,000$ system, assuming neither paging nor swapping problems arise?

**solution:**

For $\mathbf{A} \in \mathbb{C}^{m \times m}$, Gaussian elimination "costs" about $(2/3)m^3$ flops. Therefore, when neither paging nor swap are concerns, the execution time should be proportional to $m^3$. Hence, in this case, solving a $3,000 \times 3,000$ system should cost:

$$\left( \frac{3000}{1000} \right)^3 = 27$$

times as long as solving a $1,000 \times 1,000$ system, for this PC,

$$9 \text{ (2 sec)} = 54 \text{ sec}$$

3. Consider the following system:

$$\begin{array}{rcrcrcl} 2.25\ x_1 & - & 3.31\ x_2 & + & 1.28\ x_3 & = & 3.43 \\ 7.88\ x_1 & - & 11.5\ x_2 & + & 7.15\ x_3 & = & 6.35 \\ 4.23\ x_1 & + & 3.70\ x_2 & + & 2.85\ x_3 & = & 6.39 \end{array}$$

a. Simulate the solution of this system by Gaussian elimination without partial pivoting in a three-digit, decimal machine with rounding of all intermediate results.

**solution:**

Augmented Matrix:

$$\begin{bmatrix} 2.25 & -3.31 & 1.28 & 3.43 \\ 7.88 & -11.5 & 7.15 & 6.35 \\ 4.23 & 3.70 & 2.85 & 6.39 \end{bmatrix}$$

$$\begin{array}{c} \\ R_2 - (3.50)R_1 \\ R_3 - (1.88)R_1 \end{array} \quad \begin{bmatrix} 2.25 & -3.31 & 1.28 & 3.43 \\ 0 & 0.100 & 2.67 & -5.65 \\ 0 & 9.22 & 0.440 & -0.0600 \end{bmatrix}$$

$$\begin{array}{c} \\ \\ R_3 - (99.2)R_2 \end{array} \quad \begin{bmatrix} 2.25 & -3.31 & 1.28 & 3.43 \\ 0 & 0.100 & 2.67 & -5.65 \\ 0 & 0 & -265. & 560. \end{bmatrix}$$

So:

$$x_3 = \frac{560.}{-265.} = -2.11$$

$$x_2 = \frac{-5.65 - 2.67x_3}{.100} = \frac{-5.65 - \overbrace{(2.67)(-2.11)}^{-5.633\ldots}}{.100}$$
$$= \frac{-5.65 + 5.63}{.100} = \frac{-.02}{.100} = -.200$$

$$x_1 = \frac{3.43 + 3.31x_2 - 1.28x_3}{2.25} = \frac{3.43 + \overbrace{(3.31)(-.200)}^{-.662} - \overbrace{(1.28)(-2.11)}^{-2.70\ldots}}{2.25}$$
$$= \frac{3.43 - 0.66 + 2.70}{2.25} = \frac{5.47}{2.25} = 2.43$$

b. Repeat the solution of part a., but this time simulate Gaussian elimination with partial pivoting in the same machine.

**solution:**

Original Augmented Matrix is:

$$\begin{bmatrix} 2.25 & -3.31 & 1.28 & 3.43 \\ 7.88 & -11.5 & 7.15 & 6.35 \\ 4.23 & 3.70 & 2.85 & 6.39 \end{bmatrix}$$

Interchange to get largest element on the diagonal in the first column:

$$\begin{array}{c} R_1 \leftrightarrow R_2 \\ R_1 \leftrightarrow R_2 \\ \phantom{x} \end{array} \quad \begin{bmatrix} 7.88 & -11.5 & 7.15 & 6.35 \\ 2.25 & -3.31 & 1.28 & 3.43 \\ 4.23 & 3.70 & 2.85 & 6.39 \end{bmatrix}$$

Eliminate in the first column:

$$\begin{array}{c} \phantom{x} \\ R_2 - (.286)R_1 \\ R_3 - (.537)R_1 \end{array} \quad \begin{bmatrix} 7.88 & -11.5 & 7.15 & 6.35 \\ 0 & -0.0200 & -0.760 & 1.61 \\ 0 & 9.88 & -0.990 & 2.98 \end{bmatrix}$$

Interchange to get largest element on the diagonal in the second column:

$$\begin{array}{c} \phantom{x} \\ R_2 \leftrightarrow R_3 \\ R_2 \leftrightarrow R_3 \end{array} \quad \begin{bmatrix} 7.88 & -11.5 & 7.15 & 6.35 \\ 0 & 9.88 & -0.990 & 2.98 \\ 0 & -0.0200 & -0.760 & 1.61 \end{bmatrix}$$

Eliminate in the second column:

$$\begin{array}{c} \phantom{x} \\ \phantom{x} \\ R_3 - (.00202)R_2 \end{array} \quad \begin{bmatrix} 7.88 & -11.5 & 7.15 & 6.35 \\ 0 & 9.88 & -0.990 & 2.98 \\ 0 & 0 & -0.762 & 1.62 \end{bmatrix}$$

**solution:**

b. (cont) So:

$$x_3 = \frac{1.62}{-0.762} = -2.13$$

$$x_2 = \frac{2.98 + .990x_3}{9.88} = \frac{2.98 + \overbrace{(.990)(-2.13)}^{-2.1087}}{9.88}$$
$$= \frac{2.98 - 2.11}{9.88} = \frac{.87}{9.88} = .0881$$

$$x_1 = \frac{6.35 + 11.5x_2 - 7.15x_3}{7.88} = \frac{6.35 + \overbrace{(11.5)(.0881)}^{1.0132} - \overbrace{(7.15)(-2.13)}^{15.2295}}{7.88}$$
$$= \frac{6.35 + 1.01 + 15.2}{7.88} = \frac{22.6}{7.88} = 2.87$$

c. Repeat the solution of part a., but this time simulate Gaussian elimination with full pivoting in the same machine.

**solution:**

Original Augmented Matrix is:

$$\begin{bmatrix} 2.25 & -3.31 & 1.28 & 3.43 \\ 7.88 & -11.5 & 7.15 & 6.35 \\ 4.23 & 3.70 & 2.85 & 6.39 \end{bmatrix}$$

Interchange to get largest element on the diagonal in the first column:

$$\begin{array}{c} R_1 \leftrightarrow R_2 \\ R_1 \leftrightarrow R_2 \end{array} \qquad \begin{array}{c} \begin{bmatrix} -11.5 & 7.88 & 7.15 & 6.35 \\ -3.31 & 2.25 & 1.28 & 3.43 \\ 3.70 & 4.23 & 2.85 & 6.39 \end{bmatrix} \\ C_1 \leftrightarrow C_2 \end{array}$$

**solution:**

c. (cont) (Note the solution vector, in terms of the original unknowns, will now be:

$$\hat{\mathbf{x}} = [\ x_2\ ,\ \ x_1\ ,\ \ x_3\ \ ]\ \ )$$

Eliminate in the first column:

$$\begin{array}{c} \\ R_2 - (.288)R_1 \\ R_3 - (-.322)R_1 \end{array} \quad \begin{bmatrix} -11.5 & 7.88 & 7.15 & 6.35 \\ 0 & -0.0194 & -0.779 & 1.60 \\ 0 & 6.77 & 5.15 & 8.43 \end{bmatrix}$$

Interchange to get largest element on the diagonal in the second column:

$$\begin{array}{c} \\ R_2 \leftrightarrow R_3 \\ R_2 \leftrightarrow R_3 \end{array} \quad \begin{bmatrix} -11.5 & 7.88 & 7.15 & 6.35 \\ 0 & 6.77 & 5.15 & 8.43 \\ 0 & -0.0194 & -0.779 & 1.60 \end{bmatrix}$$

Eliminate in the second column:

$$\begin{array}{c} \\ \\ R_3 - (-.00287)R_2 \end{array} \quad \begin{bmatrix} -11.5 & 7.88 & 7.15 & 6.35 \\ 0 & 6.77 & 5.15 & 8.43 \\ 0 & 0 & -0.764 & 1.62 \end{bmatrix}$$

And so:

$$\hat{x}_3 = \frac{1.62}{-0.764} = -2.12$$

$$\hat{x}_2 = \frac{8.43 - 5.15\hat{x}_3}{6.77} = \frac{8.43 - \overbrace{(5.15)(-2.12)}^{-10.918}}{6.77}$$
$$= \frac{8.43 + 10.9}{6.77} = \frac{19.3}{6.77} = 2.85$$

$$\hat{x}_1 = \frac{6.35 - 7.88\hat{x}_2 - 7.15\hat{x}_3}{-11.5} = \frac{6.35 - \overbrace{(7.88)(2.85)}^{22.458} - \overbrace{(7.15)(-2.12)}^{-15.158}}{-11.5}$$
$$= \frac{6.35 - 22.5 + 15.2}{-11.5} = \frac{-.950}{-11.5} = 0.0826$$

**solution:**

c. (cont) (We would note that the numerator in the computation for $\hat{x}_1$ encountered some potentially fairly significant cancellation, and so we should not be too surprised should that value have a relatively large(r) error.)

At this point then, we have the solution to the fully pivoted system as

$$\hat{\mathbf{x}} = \begin{bmatrix} 0.0826 \\ 2.85 \\ -2.12 \end{bmatrix}$$

Recalling that the solution to the original system is related to this by:

$$\hat{\mathbf{x}} = [\ x_2\ ,\ \ x_1\ ,\ \ x_3\ \ ]\ ,$$

we finally have that the computed solution to the original system, using full pivoting, is:

$$\mathbf{x} = \begin{bmatrix} 2.85 \\ 0.0826 \\ -2.12 \end{bmatrix}$$

d. Compare your solutions from parts a., b. and c. above to the true (full MATLAB precision) solution.

**solution:**

| True (MATLAB) | Without Pivoting | Partial Pivoting | Full Pivoting |
|---|---|---|---|
| $\begin{bmatrix} 2.865178\ldots \\ 0.089165\ldots \\ -2.126181\ldots \end{bmatrix}$ | $\begin{bmatrix} 2.43 \\ -0.200 \\ -2.11 \end{bmatrix}$ | $\begin{bmatrix} 2.87 \\ 0.0881 \\ -2.13 \end{bmatrix}$ | $\begin{bmatrix} 2.85 \\ 0.0826 \\ -2.12 \end{bmatrix}$ |

Either of the three-digit solutions with pivoting are obviously far more accurate than the solution without pivoting!

e. Calculate the residuals for each of the solutions from parts a. and b. above. What can you conclude about the accuracy of the solutions from the size of their respective residuals?

**solution:**

For the solution without pivoting (using MATLAB)

$$\mathbf{r} = \begin{bmatrix} 3.43 \\ 6.35 \\ 6.39 \end{bmatrix} - \begin{bmatrix} 2.25 & -3.31 & 1.28 \\ 7.88 & -11.5 & 7.15 \\ 4.23 & 3.70 & 2.85 \end{bmatrix} \begin{bmatrix} 2.43 \\ -0.200 \\ -2.11 \end{bmatrix} = \begin{bmatrix} 0.0013 \\ -0.0119 \\ 2.8646 \end{bmatrix}$$

This residual is not bad, *except* for the third component, which is **awful**. That very strongly suggest we do not have a good solution. For the solution with partial pivoting, the residual is

$$\mathbf{r} = \begin{bmatrix} 3.43 \\ 6.35 \\ 6.39 \end{bmatrix} - \begin{bmatrix} 2.25 & -3.31 & 1.28 \\ 7.88 & -11.5 & 7.15 \\ 4.23 & 3.70 & 2.85 \end{bmatrix} \begin{bmatrix} 2.87 \\ 0.0881 \\ -2.13 \end{bmatrix} = \begin{bmatrix} -0.009489 \\ -0.022950 \\ -0.005570 \end{bmatrix}$$

which is quite acceptable. Finally, the residual for the solution with full pivoting is:

$$\mathbf{r} = \begin{bmatrix} 3.43 \\ 6.35 \\ 6.39 \end{bmatrix} - \begin{bmatrix} 2.25 & -3.31 & 1.28 \\ 7.88 & -11.5 & 7.15 \\ 4.23 & 3.70 & 2.85 \end{bmatrix} \begin{bmatrix} 2.85 \\ 0.0826 \\ -2.12 \end{bmatrix} = \begin{bmatrix} 0.004506 \\ -0.000100 \\ 0.070880 \end{bmatrix}$$

which is also quite acceptable! (Although, the value in the third coordinate is a bit larger than machine precision.) Unfortunately, as we know, the fact that the residuals for both pivoting strategies are relative small doesn't necessarily mean the solutions are any good.!)

4. Consider the following system:

$$\begin{array}{rcrcrcl} 4.55\ x_1 & + & 2.39\ x_2 & + & 4.18\ x_3 & = & 7.52 \\ 2.40\ x_1 & + & 2.04\ x_2 & + & 3.75\ x_3 & = & 4.77 \\ 3.19\ x_1 & - & 2.06\ x_2 & - & 4.23\ x_3 & = & 1.45 \end{array}$$

a. Simulate the solution of this system by Gaussian elimination without partial pivoting in a three-digit, decimal machine with rounding of all intermediate results.

**solution:**

Augmented Matrix:

$$\begin{bmatrix} 4.55 & 2.39 & 4.18 & 7.52 \\ 2.40 & 2.04 & 3.75 & 4.77 \\ 3.19 & -2.06 & -4.23 & 1.45 \end{bmatrix}$$

Eliminate first column:

$$\begin{array}{c} \\ R_2 - (.527)R_1 \\ R_3 - (.701)R_1 \end{array} \quad \begin{bmatrix} 4.55 & 2.39 & 4.18 & 7.52 \\ 0 & 0.780 & 1.55 & 0.810 \\ 0 & -3.74 & -7.16 & -3.82 \end{bmatrix}$$

Eliminate second column:

$$\begin{array}{c} \\ \\ R_3 + (4.79)R_2 \end{array} \quad \begin{bmatrix} 4.55 & 2.39 & 4.18 & 7.52 \\ 0 & 0.780 & 1.55 & 0.810 \\ 0 & 0 & 0.260 & 0.0600 \end{bmatrix}$$

Back substitution:

$$x_3 = \frac{.0600}{.260} = .231$$

$$x_2 = \frac{0.810 - 1.55x_3}{.780} = \frac{0.810 - \overbrace{(1.55)(.231)}^{0.35805}}{.780} = \frac{0.810 - 0.358}{.780} = \frac{.452}{.780} = .579$$

$$\begin{aligned} x_1 &= \frac{7.52 - 2.39x_2 - 4.18x_3}{4.55} = \frac{7.52 - \overbrace{(2.39)(.579)}^{1.38381} - \overbrace{(4.18)(.231)}^{0.96558}}{4.55} \\ &= \frac{7.52 - 1.38 - 0.966}{4.55} = \frac{5.17}{4.55} = 1.14 \end{aligned}$$

b. Repeat the solution of part a., but this time simulate Gaussian elimination with partial pivoting in the same machine.

**solution:**

Augmented Matrix:

$$\begin{bmatrix} 4.55 & 2.39 & 4.18 & 7.52 \\ 2.40 & 2.04 & 3.75 & 4.77 \\ 3.19 & -2.06 & -4.23 & 1.45 \end{bmatrix}$$

Largest element in first column is already on the diagonal, so eliminate first column:

$$\begin{array}{c} \\ R_2 - (.527)R_1 \\ R_3 - (.701)R_1 \end{array} \quad \begin{bmatrix} 4.55 & 2.39 & 4.18 & 7.52 \\ 0 & 0.780 & 1.55 & 0.810 \\ 0 & -3.74 & -7.16 & -3.82 \end{bmatrix}$$

Interchange rows to get largest element on the diagonal in the second column:

$$\begin{array}{c} \\ R_2 \longrightarrow R_3 \\ R_2 \longrightarrow R_3 \end{array} \quad \begin{bmatrix} 4.55 & 2.39 & 4.18 & 7.52 \\ 0 & -3.74 & -7.16 & -3.82 \\ 0 & 0.780 & 1.55 & 0.810 \end{bmatrix}$$

Eliminate second column:

$$\begin{array}{c} \\ \\ R_3 + (.209)R_2 \end{array} \quad \begin{bmatrix} 4.55 & 2.39 & 4.18 & 7.52 \\ 0 & -3.74 & -7.16 & -3.82 \\ 0 & 0 & 0.0500 & 0.0120 \end{bmatrix}$$

Back substitution:

$$x_3 = \frac{.0120}{.0500} = .240$$

$$\begin{aligned} x_2 = \frac{-3.82 + 7.16x_3}{-3.74} &= \frac{-3.82 - \overbrace{(7.16)(.240)}^{1.7184}}{-3.74} \\ &= \frac{-3.82 + 1.72}{-3.74} = \frac{-2.10}{-3.74} = .561 \end{aligned}$$

$$\begin{aligned} x_1 = \frac{7.52 + 2.39x_2 - 4.18x_3}{4.55} &= \frac{7.52 - \overbrace{(2.39)(.561)}^{1.34079} - \overbrace{(4.18)(.240)}^{1.0032}}{4.55} \\ &= \frac{7.52 - 1.34 - 1.00}{4.55} = \frac{5.18}{4.55} = 1.14 \end{aligned}$$

c. Compare your solutions from parts a. and b. above to the true (full MATLAB precision) solution.

**solution:**

| True (MATLAB) | Without Pivoting | With Pivoting |
|---|---|---|
| $\begin{bmatrix} 1.125590\ldots \\ 0.799300\ldots \\ 0.116802\ldots \end{bmatrix}$ | $\begin{bmatrix} 1.14 \\ 0.579 \\ 0.231 \end{bmatrix}$ | $\begin{bmatrix} 1.14 \\ 0.561 \\ 0.240 \end{bmatrix}$ |

Neither solution is really that good! (More precisely, both agree fairly well in the first component, but their second and third components don't even have a single digit correct!

d. Calculate the residuals for each of the solutions from parts a. and b. above. What can you conclude about the accuracy of the solutions from the size of their respective residuals?

**solution:**

For the solution without pivoting (using MATLAB)

$$\mathbf{r} = \begin{bmatrix} 7.52 \\ 4.77 \\ 1.45 \end{bmatrix} - \begin{bmatrix} 4.55 & 2.39 & 4.18 \\ 2.40 & 2.04 & 3.75 \\ 3.19 & -2.06 & -4.23 \end{bmatrix} \begin{bmatrix} 1.14 \\ 0.579 \\ 0.231 \end{bmatrix} = \begin{bmatrix} -0.01639 \\ -0.01341 \\ -0.01673 \end{bmatrix}$$

This residual is not bad. For the solution with pivoting, the residual is

$$\mathbf{r} = \begin{bmatrix} 7.52 \\ 4.77 \\ 1.45 \end{bmatrix} - \begin{bmatrix} 4.55 & 2.39 & 4.18 \\ 2.40 & 2.04 & 3.75 \\ 3.19 & -2.06 & -4.23 \end{bmatrix} \begin{bmatrix} 1.14 \\ 0.561 \\ 0.240 \end{bmatrix} = \begin{bmatrix} -0.01099 \\ -0.01044 \\ -0.01574 \end{bmatrix}$$

Which is also quite acceptable! Unfortunately, as we've already noted, small residuals doesn't necessarily mean the solution is any good! And, in this example, that is precisely the case! )

5. Calculate the true (MATLAB) solution for the system:

$$\begin{array}{rcrcrcl} 4.55\ x_1 & + & 2.39\ x_2 & + & 4.18\ x_3 & = & 7.56 \\ 2.40\ x_1 & + & 2.04\ x_2 & + & 3.75\ x_3 & = & 4.72 \\ 3.19\ x_1 & - & 2.06\ x_2 & - & 4.23\ x_3 & = & 1.50 \end{array}$$

(Notice the left-hand side here is identical to, and the right-hand side only a relatively small perturbation of problem 4.) Compare the exact (MATLAB) solutions to both problems. What does that comparison suggest about the condition of this matrix?

**solution:**

Using MATLAB, the calculated solution to this system is

$$\mathbf{x} = \begin{bmatrix} 1.023110\ldots \\ 3.279168\ldots \\ -1.179991\ldots \end{bmatrix}$$

while the solution to the system in problem 4 is:

$$\begin{bmatrix} 1.125590\ldots \\ 0.799300\ldots \\ 0.116802\ldots \end{bmatrix}$$

Observe these solutions are drastically different, especially in the second and third components, even thought the came from the same basic system with only slightly changed right-hand sides. This is classically characteristic behavior for so-called *ill-conditioned* systems.

6. Consider the matrices:

$$\mathbf{P} = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \end{bmatrix} \qquad \text{and} \qquad \mathbf{A} = \begin{bmatrix} 3 & -1 & 2 & 4 \\ -4 & 2 & -3 & -3 \\ 2 & -2 & 5 & 2 \\ 3 & -3 & 1 & -1 \end{bmatrix}$$

a. Show that $\mathbf{P}$ is a permutation matrix by showing that it can be created as a product of elementary permutation matrices.

**solution:**

Multiplication on the left by $\mathbf{P}$ should produce:

$R_3$ replaces $R_1$ , $R_1$ replaces $R_2$ , $R_4$ replaces $R_3$ , and $R_2$ replaces $R_4$

We can accomplish this by performing, in sequence:
(1) Interchanging rows $R_1$ and $R_3$ (i.e. $R_1 \leftrightarrow R_3$)
(2) Interchanging rows $R_2$ and $R_3$ (the old $R_1$) (i.e. $R_2 \leftrightarrow R_3$)
(3) Interchanging rows $R_3$ (the old $R_2$) and $R_4$ (i.e. $R_3 \leftrightarrow R_4$)

This is equivalent to the product:

$$\underbrace{\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}}_{\mathbf{P}^{(3)}} \underbrace{\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}}_{\mathbf{P}^{(2)}} \underbrace{\begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}}_{\mathbf{P}^{(1)}} = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \end{bmatrix} \equiv \mathbf{P}$$

Note the order of the terms in the product must be as shown in order to preroms step (1) first.

b. Show that $\mathbf{P}^{(i)}\,\mathbf{P}^{(i)} = \mathbf{I}$, where $\mathbf{P}^{(i)}$ denotes any one of the elementary permutation matrices you found in part a., and therefore that ${\mathbf{P}^{(i)}}^{-1} = \mathbf{P}^{(i)}$. Why should this be true in general for *elementary* permutations? Why, however, should $\mathbf{P}^{-1} = \mathbf{P}$ **not** be true for more general permutations, e.g. why does $\mathbf{P}\,\mathbf{P} \neq \mathbf{I}$ here?

**solution:**

Direct MATLAB calculation shows:

$$\mathbf{P}^{(1)}\,\mathbf{P}^{(1)} = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \equiv \mathbf{I}$$

Similar computations verify the same is true for $\mathbf{P}^{(2)}$ and $\mathbf{P}^{(3)}$. This should occur because all of these are *elementary* permutations, i.e. corresponding to a *single* row interchange, e.g. $R_1 \leftrightarrow R_3$. Clearly, if we perform the same exact interchange twice in succession on the same matrix, e.g. interchange the first and third rows, then again interchange the first (old third) and third(old first) rows, we will end up exactly where we started! In other words, we will have the original matrix, which is precised the result of simply multiplying by the identity.

However, for

$$\mathbf{P} = \mathbf{P}^{(3)}\ \mathbf{P}^{(2)}\ \mathbf{P}^{(1)}$$

elementary matrix properties imply that

$$\mathbf{P}^{-1} = \mathbf{P}^{(1)^{-1}}\ \mathbf{P}^{(2)^{-1}}\ \mathbf{P}^{(3)^{-1}} = \mathbf{P}^{(1)}\ \mathbf{P}^{(2)}\ \mathbf{P}^{(3)}$$

which is not equal to $\mathbf{P}$ unless these permutations *commute*, which, of course, matrix multiplications generally do not.

You might also note, that in general, **elementary** permutation matrices are symmetric, i.e.

$$\mathbf{P}^{(i)^T} = \mathbf{P}^{(i)} = \mathbf{P}^{(i)^{-1}}$$

and therefore, in this case (and actually in general)

$$\mathbf{P}^{-1} = \mathbf{P}^{(1)}\mathbf{P}^{(2)}\mathbf{P}^{(3)} = \mathbf{P}^{(1)^T}\mathbf{P}^{(2)^T}\mathbf{P}^{(3)^T} = \left(\mathbf{P}^{(3)}\mathbf{P}^{(2)}\mathbf{P}^{(1)}\right)^T = \mathbf{P}^T$$

c. Compute $\mathbf{P\,A}$ and $\mathbf{A\,P}$ and show that the results are as predicted by theory.

**solution:**

Theory predicts that multipying a matrix $\mathbf{A}$ on the left by a permutation matrix will produce the corresponding permutation of the **rows** of $\mathbf{A}$, while multiplying on the right by the same permutation will produce the corresponding permutation of the **columns** of $\mathbf{A}$. MATLAB shows:

$$\mathbf{P\,A} = \begin{array}{cl} \begin{bmatrix} 2 & -2 & 5 & 2 \\ 3 & -1 & 2 & 4 \\ 3 & -3 & 1 & -1 \\ -4 & 2 & -3 & -3 \end{bmatrix} & \begin{array}{l} \longleftarrow \ \text{Original third row} \\ \longleftarrow \ \text{Original first row} \\ \longleftarrow \ \text{Original fourth row} \\ \longleftarrow \ \text{Original second row} \end{array} \end{array}$$

and

$$\mathbf{A\,P} = \begin{array}{l} \begin{bmatrix} -1 & 4 & 3 & 2 \\ 2 & -3 & -4 & -3 \\ -2 & 2 & 2 & 5 \\ -3 & -1 & 3 & 1 \end{bmatrix} \\ \begin{array}{cccc} \uparrow & \uparrow & \uparrow & \uparrow \\ (2) & (4) & (1) & (2) \end{array} \longleftarrow \text{Original column number} \end{array}$$

7. Solve the following system by **L U** Decomposition with Partial Pivoting, and forward/backward substitution:

$$\begin{array}{rcrcrcr} 3x_1 & - & x_2 & + & 2x_3 & = & 1 \\ 6x_1 & - & 2x_2 & + & 5x_3 & = & -1 \\ -3x_1 & + & 2x_2 & - & x_3 & = & -5 \end{array}$$

**solution:**

Original Matrices:

$$\mathbf{U}_{work} = \begin{bmatrix} 3 & -1 & 2 \\ 6 & -2 & 5 \\ -3 & 2 & -1 \end{bmatrix}, \ \mathbf{L}_{work} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \ \mathbf{p}_{work} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$$

Interchange first and second rows to place largest element on diagonal, i.e.:
(i) $R_1 \leftrightarrow R_2$ to yeild:

$$\mathbf{U}_{work} = \begin{bmatrix} 6 & -2 & 5 \\ 3 & -1 & 2 \\ -3 & 2 & -1 \end{bmatrix}, \ \mathbf{L}_{work} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \ \mathbf{p}_{work} = \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix}$$

(Note there are no subdiagonal elements in **L** yet.): Then eliminate in first column, i.e.:
(i) $R_2 - \frac{1}{2}R_1$
(ii) $R_3 - (-\frac{1}{2})R_1$

to yield

$$\mathbf{U}_{work} = \begin{bmatrix} 6 & -2 & 5 \\ 0 & 0 & -\frac{1}{2} \\ 0 & 1 & \frac{3}{2} \end{bmatrix}, \ \mathbf{L}_{work} = \begin{bmatrix} 1 & 0 & 0 \\ \frac{1}{2} & 1 & 0 \\ -\frac{1}{2} & 0 & 1 \end{bmatrix}, \ \mathbf{p}_{work} = \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix}$$

Zero on diagonal in second column mandates row interchange. (Interchange subdiagonal elements of **L** only!) Perform
(i) $R_2 \leftrightarrow R_3$ yielding:

$$\mathbf{U}_{work} = \begin{bmatrix} 6 & -2 & 5 \\ 0 & 1 & \frac{3}{2} \\ 0 & 0 & -\frac{1}{2} \end{bmatrix}, \ \mathbf{L}_{work} = \begin{bmatrix} 1 & 0 & 0 \\ -\frac{1}{2} & 1 & 0 \\ \frac{1}{2} & 0 & 1 \end{bmatrix}, \ \mathbf{p}_{work} = \begin{bmatrix} 2 \\ 3 \\ 1 \end{bmatrix}$$

**solution:**

Factorization complete - $\mathbf{P}\,\mathbf{A} = \mathbf{L}\,\mathbf{U}$ , where

$$\mathbf{P} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix} \ , \ \mathbf{L} = \begin{bmatrix} 1 & 0 & 0 \\ -\frac{1}{2} & 1 & 0 \\ \frac{1}{2} & 0 & 1 \end{bmatrix} \ , \ \mathbf{U} = \begin{bmatrix} 6 & -2 & 5 \\ 0 & 1 & \frac{3}{2} \\ 0 & 0 & -\frac{1}{2} \end{bmatrix}$$

Solve original system as:

$$\mathbf{L}\,\mathbf{z} = \mathbf{P}\,\mathbf{b}$$
$$\mathbf{U}\,\mathbf{x} = \mathbf{z}$$

Forward substitution:

$$\begin{array}{rcccll} z_1 & & & & & = -1 \\ -\frac{1}{2}z_1 & + & z_2 & & & = -5 \\ \frac{1}{2}z_1 & & & + & z_3 & = \ \ 1 \end{array}$$

implies

$$z_1 = -1$$

$$z_2 = -5 + \frac{1}{2}z_1 = -5 + (\frac{1}{2})(-1) = -\frac{11}{5}$$

$$z_3 = 1 - \frac{1}{2}z_1 = 1 - (\frac{1}{2})(-1) = \frac{3}{2}$$

Backward substitution:

$$\begin{array}{rcrcrcc} 6x_1 & - & 2x_2 & + & 5x_3 & = & -1 \\ & & x_2 & + & \frac{3}{2}x_3 & = & -\frac{11}{5} \\ & & & - & \frac{1}{2}x_3 & = & \frac{3}{2} \end{array}$$

implies

$$x_3 = \frac{3/2}{-1/2} = -3$$

$$x_2 = -\frac{11}{2} - \frac{3}{2}x_3 = -\frac{11}{2} - \left(\frac{3}{2}\right)(-3) = -\frac{2}{2} = -1$$

$$x_1 = \frac{-1 + 2x_2 - 5x_3}{6} = \frac{-1 + (2)(-1) - (5)(-3)}{6} = \frac{12}{6} = 2$$

8. Solve the following system by $\mathbf{L}\,\mathbf{U}$ Decomposition with Partial Pivoting, and forward/backward substitution:

$$\begin{array}{rcrcrcrcr} 2x_1 & + & x_2 & - & 2x_3 & - & 2x_4 & = & 2 \\ 3x_1 & + & x_2 & + & 6x_3 & - & 3x_4 & = & -6 \\ -2x_1 & + & x_2 & + & x_3 & + & 2x_4 & = & -1 \\ -x_1 & + & x_2 & - & 8x_3 & + & 3x_4 & = & 10 \end{array}$$

**solution:**

Original Matrices:

$$\mathbf{U}_{work} = \begin{bmatrix} 2 & 1 & -2 & -2 \\ 3 & 1 & 6 & -3 \\ -2 & 1 & 1 & 2 \\ -1 & 1 & -8 & 3 \end{bmatrix} , \; \mathbf{L}_{work} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} , \; \mathbf{p}_{work} = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix}$$

Interchanging first and second rows to place largest element on diagonal, i.e:

(i) $R_1 \leftrightarrow R_2$

yields:

$$\mathbf{U}_{work} = \begin{bmatrix} 3 & 1 & 6 & -3 \\ 2 & 1 & -2 & -2 \\ -2 & 1 & 1 & 2 \\ -1 & 1 & -8 & 3 \end{bmatrix} , \; \mathbf{L}_{work} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} , \; \mathbf{p}_{work} = \begin{bmatrix} 2 \\ 1 \\ 3 \\ 4 \end{bmatrix}$$

(Note there are no subdiagonal elements in $\mathbf{L}_{work}$ yet!) Then eliminate in first column, i.e.:

(i) $R_2 - (\frac{2}{3})R_1$

(ii) $R_3 - (-\frac{2}{3})R_1$

(iii) $R_4 - (-\frac{1}{3})R_1$

to yield:

$$\mathbf{U}_{work} = \begin{bmatrix} 3 & 1 & 6 & -3 \\ 0 & \frac{1}{3} & -6 & 0 \\ 0 & \frac{5}{3} & 5 & 0 \\ 0 & \frac{4}{3} & -6 & 2 \end{bmatrix} , \; \mathbf{L}_{work} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ \frac{2}{3} & 1 & 0 & 0 \\ -\frac{2}{3} & 0 & 1 & 0 \\ -\frac{1}{3} & 0 & 0 & 1 \end{bmatrix} , \; \mathbf{p}_{work} = \begin{bmatrix} 2 \\ 1 \\ 3 \\ 4 \end{bmatrix}$$

**solution:**

To place the largest element in the pivot in the second column, we must now interchange

(i) $\mathrm{R}_2 \leftrightarrow \mathrm{R}_3$

in $\mathbf{U}_{work}$ and $\mathbf{p}_{work}$, and in the subdiagonal portion of $\mathbf{L}_{work}$:

$$\mathbf{U}_{work} = \begin{bmatrix} 3 & 1 & 6 & -3 \\ 0 & \frac{5}{3} & 5 & 0 \\ 0 & \frac{1}{3} & -6 & 0 \\ 0 & \frac{4}{3} & -6 & 2 \end{bmatrix} , \ \mathbf{L}_{work} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -\frac{2}{3} & 1 & 0 & 0 \\ \frac{2}{3} & 0 & 1 & 0 \\ -\frac{1}{3} & 0 & 0 & 1 \end{bmatrix} , \ \mathbf{p}_{work} = \begin{bmatrix} 2 \\ 3 \\ 1 \\ 4 \end{bmatrix}$$

Next eliminate in the second column with:

(i) $R_3 - (\frac{1}{5})R_2$

(ii) $R_4 - (\frac{4}{5})R_2$

yielding:

$$\mathbf{U}_{work} = \begin{bmatrix} 3 & 1 & 6 & -3 \\ 0 & \frac{5}{3} & 5 & 0 \\ 0 & 0 & -7 & 0 \\ 0 & 0 & -10 & 2 \end{bmatrix} , \ \mathbf{L}_{work} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -\frac{2}{3} & 1 & 0 & 0 \\ \frac{2}{3} & \frac{1}{5} & 1 & 0 \\ -\frac{1}{3} & \frac{4}{5} & 0 & 1 \end{bmatrix} , \ \mathbf{p}_{work} = \begin{bmatrix} 2 \\ 3 \\ 1 \\ 4 \end{bmatrix}$$

To place the largest element in the pivot in the third column, we must now interchange

(i) $\mathrm{R}_3 \leftrightarrow \mathrm{R}_4$

in $\mathbf{U}_{work}$ and $\mathbf{p}_{work}$, and in the subdiagonal portion of $\mathbf{L}_{work}$:

$$\mathbf{U}_{work} = \begin{bmatrix} 3 & 1 & 6 & -3 \\ 0 & \frac{5}{3} & 5 & 0 \\ 0 & 0 & -10 & 2 \\ 0 & 0 & -7 & 0 \end{bmatrix} , \ \mathbf{L}_{work} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -\frac{2}{3} & 1 & 0 & 0 \\ -\frac{1}{3} & \frac{4}{5} & 1 & 0 \\ \frac{2}{3} & \frac{1}{5} & 0 & 1 \end{bmatrix} , \ \mathbf{p}_{work} = \begin{bmatrix} 2 \\ 3 \\ 4 \\ 1 \end{bmatrix}$$

Finally we elimiante in the third column: (i) $\mathrm{R}_4 - (\frac{7}{10})\mathrm{R}_3$

yielding:

$$\mathbf{U}_{work} = \begin{bmatrix} 3 & 1 & 6 & -3 \\ 0 & \frac{5}{3} & 5 & 0 \\ 0 & 0 & -10 & 2 \\ 0 & 0 & 0 & -\frac{7}{5} \end{bmatrix} , \ \mathbf{L}_{work} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -\frac{2}{3} & 1 & 0 & 0 \\ -\frac{1}{3} & \frac{4}{5} & 1 & 0 \\ \frac{2}{3} & \frac{1}{5} & \frac{7}{10} & 1 \end{bmatrix} ,$$

**solution:**

and

$$\mathbf{p}_{work} = \begin{bmatrix} 2 \\ 3 \\ 4 \\ 1 \end{bmatrix} \quad \Longrightarrow \quad \mathbf{P} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{bmatrix}$$

We now solve

$$\mathbf{A}\,\mathbf{x} = \mathbf{b} \qquad \text{using} \qquad \mathbf{P}\,\mathbf{A} = \mathbf{L}\,\mathbf{U}$$

by forward and backward substitution as:

$$\begin{aligned} \mathbf{L}\,\mathbf{z} &= \mathbf{P}\,\mathbf{b} \\ \mathbf{U}\,\mathbf{x} &= \mathbf{z} \end{aligned}$$

Forward substitution:

$$\begin{array}{rcrcrcrcr} z_1 & & & & & & & = & -6 \\ -\frac{2}{3}z_1 & + & z_2 & & & & & = & -1 \\ -\frac{1}{3}z_1 & + & \frac{4}{5}z_2 & + & z_3 & & & = & 10 \\ \frac{2}{3}z_1 & + & \frac{1}{5}z_2 & + & \frac{7}{10}z_3 & + & z_4 & = & 2 \end{array}$$

yields

$$z_1 = -6$$

$$z_2 = -1 + \frac{2}{3}z_1 = -1 + (\frac{2}{3})(-6) = -5$$

$$z_3 = 10 + \frac{1}{3}z_1 - \frac{4}{5}z_2 = 10 + (\frac{1}{3})(-6) - (\frac{4}{5})(-5) = 12$$

$$z_4 = 2 - \frac{2}{3}z_1 - \frac{1}{5}z_2 + \frac{7}{10}z_3 = 2 - (\frac{2}{3})(-6) - (\frac{1}{5})(-5) + (\frac{7}{10})(12) = -\frac{7}{5}$$

Backward substitution:

$$\begin{array}{rcrcrcrcr} 3x_1 & + & x_2 & + & 6x_3 & - & 3x_4 & = & -6 \\ & & \frac{5}{3}x_2 & + & 5x_3 & & & = & -5 \\ & & & - & 10x_3 & + & 2x_4 & = & 12 \\ & & & & & - & \frac{7}{5}x_4 & = & -\frac{7}{5} \end{array}$$

**solution:**

yields:

$$x_4 = \frac{-7/5}{-7/5} = 1$$

$$x_3 = \frac{12 - 2x_3}{-10} = \frac{12 - (2)(1)}{-10} = -1$$

$$x_2 = \frac{-5 - 5x_3}{5/3} = \frac{-5 - (5)(-1)}{5/3} = 0$$

$$x_1 = \frac{-6 - x_2 - 6x_3 + 3x_4}{3} = \frac{-6 - (0) - (6)(-1) + (3)(1)}{3} = 1$$

9. a. Show that for any invertible matrices $\mathbf{A}$ and $\mathbf{C}$ in $\mathbb{C}^{m\times m}$,

$$\begin{bmatrix} \mathbf{A} & \vdots & \mathbf{B} \\ . & \cdots & . \\ \mathbf{0} & \vdots & \mathbf{C} \end{bmatrix}^{-1} = \begin{bmatrix} \mathbf{A}^{-1} & \vdots & -\mathbf{A}^{-1}\mathbf{B}\mathbf{C}^{-1} \\ \cdots\cdot & \cdots & \cdots\cdots\cdots\cdots \\ \mathbf{0} & \vdots & \mathbf{C}^{-1} \end{bmatrix}$$

(Note that from this we can immediately conclude that the inverse of any upper triangular matrix $\mathbf{U}$ is also upper triangular, provided all the diagonal elements of $\mathbf{U}$ are non-zero.)

**solution:**

Direct computation shows

$$\begin{bmatrix} \mathbf{A} & \vdots & \mathbf{B} \\ . & \cdots & . \\ \mathbf{0} & \vdots & \mathbf{C} \end{bmatrix} \begin{bmatrix} \mathbf{A}^{-1} & \vdots & -\mathbf{A}^{-1}\mathbf{B}\mathbf{C}^{-1} \\ \cdots\cdot & \cdots & \cdots\cdots\cdots\cdots \\ \mathbf{0} & \vdots & \mathbf{C}^{-1} \end{bmatrix}$$

$$= \begin{bmatrix} \mathbf{A}\mathbf{A}^{-1} & \vdots & -\mathbf{A}\mathbf{A}^{-1}\mathbf{B}\mathbf{C}^{-1} + \mathbf{B}\mathbf{C}^{-1} \\ \cdots\cdots & \cdots & \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ \mathbf{0} & \vdots & \mathbf{C}\mathbf{C}^{-1} \end{bmatrix} =$$

$$= \begin{bmatrix} \mathbf{I} & \vdots & -\mathbf{B}\mathbf{C}^{-1} + \mathbf{B}\mathbf{C}^{-1} \\ . & \cdots & \cdots\cdots\cdots\cdots\cdots \\ \mathbf{0} & \vdots & \mathbf{I} \end{bmatrix} = \begin{bmatrix} \mathbf{I} & \vdots & \mathbf{0} \\ . & \cdots & . \\ \mathbf{0} & \vdots & \mathbf{I} \end{bmatrix}$$

and we're done.

b. Using the result from part a. above, show that, if $\mathbf{U}$ is invertible, then the diagonal elements of $\mathbf{U}^{-1}$ are precisely the reciprocals of the diagonal elements of $\mathbf{U}$.

**solution:**

Observe that the diagonal elements of $\mathbf{U}^{-1}\mathbf{U}$ are precisely the inner products of the rows of $\mathbf{U}^{-1}$ with the columns of $\mathbf{U}$. Therefore, since these diagonal elements of the product are precisely the diagonal elements of the identity,

$$1 = \sum_{k=1}^{m} \left(\mathbf{U}^{-1}\right)_{ik} u_{ki}$$

**solution:**

But, because $\mathbf{U}$ is upper triangular, then $u_{ki} = 0$ , $k > i$, and because $\mathbf{U}^{-1}$ is also upper triangular as well, then also $\left(\mathbf{U}^{-1}\right)_{ik} = 0$ , $k < i$. But this means

$$1 = \sum_{k=1}^{m} \left(\mathbf{U}^{-1}\right)_{ik} u_{ki} = \left(\mathbf{U}^{-1}\right)_{ii} u_{ii}$$

and therefore it follows immediately that

$$\left(\mathbf{U}^{-1}\right)_{ii} = \frac{1}{u_{ii}}$$

c. Based on your result from part b., show that the condition number of an upper triangular matrix $\mathbf{U}$ satisfies

$$\kappa(U) \geq \frac{\max_{i,j} |u_{ij}|}{\max_i |u_{ii}|}$$

This result, of course, explains why either a large growth factor ($\rho$) or the appearance of unavoidable small pivots is a sure sign of trouble in Gaussian elimination.

**solution:**

Recall that, for any matrix $\mathbf{A}$,

$$\| \mathbf{A} \| = \max_{\| x \| = 1} \| \mathbf{A}\mathbf{x} \|$$

Suppose now, that for some fixed values the largest element of $\mathbf{A}$ in magnitude occurs in column $J$. Now let $\hat{\mathbf{x}}$ a vector of all zeros except for $\hat{x}_J = 1$. Then, obviously

$$\mathbf{A}\hat{\mathbf{x}} = \begin{bmatrix} a_{1J} \\ a_{2J} \\ \vdots \\ a_{mJ} \end{bmatrix} \quad \Longrightarrow \quad \| A \| \geq \max_i |a_{iJ}| \equiv \max_{i,j} |a_{ij}|$$

regardless of which norm we choose to use.

**solution:**

But, based on this last observation and the result from part b., it follows immediately that

$$\| \mathbf{U} \| \geq \max_{i,j} |u_{ij}|$$

and

$$\| \mathbf{U}^{-1} \| \geq \max_{i,j} \left| \left(\mathbf{U}^{-1}\right)_{ij} \right| \geq \max_{i} \left| \left(\mathbf{U}^{-1}\right)_{ii} \right| = \max_{i} \left| \frac{1}{u_{ii}} \right| = \frac{1}{\min_i |u_{ii}|}$$

But then, by definition

$$\kappa(\mathbf{U}) = \| \mathbf{U} \| \cdot \| \mathbf{U}^{-1} \| \geq \max_{i,j} |u_{ij}| \cdot \frac{1}{\min_i |u_{ii}|} = \frac{\max_{i,j} |u_{ij}|}{\min_i |u_{ii}|}$$

10. Consider the system of linear equations:

$$\mathbf{A}\,\mathbf{x} = \mathbf{b}$$

where

$$\mathbf{A} = \begin{bmatrix} 3.44 & 12.4 \\ -0.345 & -1.32 \end{bmatrix} = \begin{bmatrix} 1.00 & 0 \\ -.100 & 1.00 \end{bmatrix} \begin{bmatrix} 3.44 & 12.4 \\ 0 & -0.0800 \end{bmatrix}$$

and

$$\mathbf{b} = \begin{bmatrix} 1.14 \\ -0.620 \end{bmatrix} \quad .$$

In a computer for which (normal) single precision is three decimal digits, with rounding of all intermediate terms, a computed solution to this problem is

$$\tilde{\mathbf{x}}^{(0)} = \begin{bmatrix} -22.5 \\ 6.33 \end{bmatrix} \quad .$$

a. Perform a single iteration of iterative improvement on the above solution.

**solution:**

First, compute the residual (using the original matrix **A** and at least six digit arithmetic):

$$\mathbf{r}^{(0)} = \mathbf{b} - \mathbf{A}\,\tilde{\mathbf{x}}^{(0)} = \begin{bmatrix} 1.14 \\ -0.620 \end{bmatrix} - \begin{bmatrix} 3.44 & 12.4 \\ -0.345 & -1.32 \end{bmatrix} \begin{bmatrix} -22.5 \\ 6.33 \end{bmatrix}$$
$$= \begin{bmatrix} 0.0480000 \\ -0.0269000 \end{bmatrix} = \begin{bmatrix} 0.0480 \\ -0.0269 \end{bmatrix}$$

(Note that after the residual has been computed in extended precision, it's perfectly permissable to drop it back to single precision. Note also that:

$$\frac{\| \mathbf{r}^{(0)} \|_\infty}{\| \mathbf{b} \|_\infty} = \frac{.048}{1.14} \doteq .042$$

which is about an order of magnitude larger than machine precision ($eps = .005$) for this machine. So we might expect this solution isn't all that good!) Next, proceed to solve:

$$\mathbf{A}\,\mathbf{e}^{(0)} = \mathbf{L}\,\mathbf{U}\,\mathbf{e}^{(0)} = \mathbf{r}^{(0)}$$

**solution:**

First forward solve:

$$\begin{bmatrix} 1.00 & 0 \\ -.100 & 1.00 \end{bmatrix} \mathbf{z}^{(0)} = \mathbf{r}^{(0)} \quad \Longrightarrow \quad \begin{array}{rcrcr} z_1^{(0)} & & & = & 0.0480 \\ -.100 z_1^{(0)} & + & z_2^{(0)} & = & -0.0269 \end{array}$$

$$\Longrightarrow \quad \begin{array}{rcccr} z_1^{(0)} & = & 0.0480 & & \\ z_2^{(0)} & = & -0.0269 - (-.100)(.0480) & = & -0.0221 \end{array}$$

Then backward solve:

$$\begin{bmatrix} 3.44 & 12.4 \\ 0 & -0.0800 \end{bmatrix} \mathbf{e}^{(0)} = \mathbf{z}^{(0)} \Longrightarrow \quad \begin{array}{rrcr} 3.44 e_1^{(0)} & +12.4 e_2^{(0)} & = & 0.0480 \\ & -0.0800 e_2^{(0)} & = & -0.0221 \end{array}$$

$$\Longrightarrow \quad \begin{array}{rcccr} e_2^{(0)} & = & \frac{-0.0221}{-0.0800} = .276 & & \\ e_1^{(0)} & = & \frac{.0480-(12.4)(.276)}{3.44} & = & -.980 \end{array}$$

or

$$\tilde{\mathbf{e}}^{(0)} = \begin{bmatrix} -0.980 \\ 0.276 \end{bmatrix}$$

Therefore:

$$\tilde{\mathbf{x}}^{(1)} = \tilde{\mathbf{x}}^{(0)} + \tilde{\mathbf{e}}^{(0)} = \begin{bmatrix} -22.5 \\ 6.33 \end{bmatrix} + \begin{bmatrix} -0.980 \\ 0.276 \end{bmatrix} = \begin{bmatrix} -23.5 \\ 6.61 \end{bmatrix}$$

b. Based on your answer to part a, do you feel this matrix is ill-conditioned in a three decimal digit machine. (*Justify your answer.*)

**solution:**

Note that

$$\frac{\| \tilde{\mathbf{e}}^{(0)} \|_\infty}{\| \tilde{\mathbf{x}}^{()} \|_\infty} = \frac{.980}{22.5} \doteq .044$$

**solution:**

This suggests that the original computation error in $\tilde{\mathbf{x}}^{()}$ was about five percent. That's not too bad, but its also quite about an order of magnitude above machine precision. Also, note the rather small pivot element in the $(2,2)$ position of $\mathbf{U}$, relative to the corresponding number in the same position of $\mathbf{A}$. Therefore, this matrix is probably at least mildly ill-conditioned for this machine. (But it's also likely not so ill-conditioned that a couple of iterations of iterative improvement would not be able to recover an accurate solution.)

We could actually check this by performing another iteration of iterative improvement. This would produce (skipping the details):

$$\mathbf{r}^{(1)} = \begin{bmatrix} 0.0160 \\ -0.00230 \end{bmatrix}$$

$$\mathbf{z}^{(1)} = \begin{bmatrix} 0.0160 \\ -0.000700 \end{bmatrix}$$

$$\tilde{\mathbf{e}}^{(1)} = \begin{bmatrix} -0.0270 \\ 0.00875 \end{bmatrix}$$

and finally

$$\tilde{\mathbf{x}}^{(2)} = \begin{bmatrix} -23.5 \\ 6.62 \end{bmatrix}$$

This solution agrees with the true (MATLAB) solution

$$\mathbf{x} = \begin{bmatrix} -23.5281\ldots \\ 6.6191\ldots \end{bmatrix}$$

to (three-digit) machine precision.

11. Find the Cholesky factorization of the symmetric, positive definite matrix

$$\mathbf{A} = \begin{bmatrix} 16 & -4 & -12 \\ -4 & 82 & -24 \\ -12 & -24 & 22 \end{bmatrix}$$

**solution:**

Cholesky factorization is a modification of Gaussian elimination, which exploits the fact that pivoting is not required with Hermitian, positive-definite matrices. (Real, symmetric matrices are automatically Hermitian!) The elimination is based on the fact that the formulas

$$\begin{bmatrix} 1/\sqrt{a_{11}} & \vdots & \mathbf{w}^H \\ \cdots\cdots & \cdots & \cdots \\ -\mathbf{w}/a_{11} & \vdots & \mathbf{I} \end{bmatrix} \begin{bmatrix} a_{11} & \vdots & \mathbf{w}^H \\ \cdots & \cdots & \cdots \\ \mathbf{w} & \vdots & \mathbf{K} \end{bmatrix} = \begin{bmatrix} \sqrt{a_{11}} & \vdots & \mathbf{w}^H/a_{11} \\ \cdots & \cdots & \cdots\cdots \\ \mathbf{0} & \vdots & \mathbf{K} - \frac{\mathbf{w}\mathbf{w}^H}{a_{11}} \end{bmatrix}$$

and

$$\mathbf{A} = \begin{bmatrix} a_{11} & \vdots & \mathbf{w}^H \\ \cdots & \cdots & \cdots \\ \mathbf{w} & \vdots & \mathbf{K} \end{bmatrix} = \begin{bmatrix} \sqrt{a_{11}} & \vdots & \mathbf{0} \\ \cdots\cdots & \cdots & \cdot \\ \mathbf{w}/\sqrt{a_{11}} & \vdots & \mathbf{I} \end{bmatrix} \begin{bmatrix} \sqrt{a_{11}} & \vdots & \mathbf{w}^H/\sqrt{a_{11}} \\ \cdots & \cdots & \cdots\cdots \\ \mathbf{0} & \vdots & \mathbf{K} - \frac{\mathbf{w}\,\mathbf{w}^H}{a_{11}} \end{bmatrix}$$

are equivalent, and therefore, we only need to store (and compute) the elements on and above the diagonal. Repeating this leads to an upper triangular matrix $\mathbf{R}$, such that

$$\mathbf{A} = \mathbf{R}^H\mathbf{R}$$

To align this problem with the above notation, consider our original matrix to be:

$$\mathbf{A} = \begin{bmatrix} 16 & \vdots & -4 & -12 \\ \cdots\cdots & \cdots & \cdots\cdots & \cdots\cdots \\ -4 & \vdots & 82 & -24 \\ -12 & & -24 & 22 \end{bmatrix}$$

and so

$$\mathbf{w} = \begin{bmatrix} -4 \\ -12 \end{bmatrix} \quad \text{and} \quad \mathbf{K} = \begin{bmatrix} 82 & -24 \\ -24 & 22 \end{bmatrix}$$

**solution:**

But now,

$$\mathbf{K} - \frac{\mathbf{w}\,\mathbf{w}^H}{a_{11}} = \begin{bmatrix} 82 & -24 \\ -24 & 22 \end{bmatrix} - \frac{1}{4}\begin{bmatrix} 16 & 48 \\ 48 & 144 \end{bmatrix}$$

which implies that after eliminating the first column, we will be left with

$$\begin{bmatrix} 4 & -1 & -3 \\ 0 & 81 & -27 \\ 0 & -27 & 13 \end{bmatrix}$$

If we now repeat the same thing on the lower right hand block here, i.e. use

$$\hat{\mathbf{K}} = \begin{bmatrix} 81 & -27 \\ -27 & 13 \end{bmatrix} \quad \Longrightarrow \quad \hat{\mathbf{w}} = [\ -27\ ] \quad \text{and} \quad \hat{\hat{\mathbf{K}}} = [\ 13\ ]$$

then Cholesky elimination on this produces

$$\hat{\hat{\mathbf{K}}} - \frac{\hat{\mathbf{w}}\,\hat{\mathbf{w}}^H}{a_{22}} = [\ 13\ ] - \frac{[\ -27\ ][\ -27\ ]}{81} = [\ 4\ ]$$

ando so eliminating on the reduced matrix $\hat{\mathbf{K}}$ will produce,

$$\begin{bmatrix} 9 & -3 \\ 0 & 4 \end{bmatrix}$$

or, when updated in the full matrix, eliminating in the second column will produce

$$\begin{bmatrix} 4 & -1 & -3 \\ 0 & 9 & -3 \\ 0 & 0 & 4 \end{bmatrix}$$

Finally, although we do not need to eliminate below the diagonal in the last column, we do need to replace the diagonal element there by its square root, i.e.

$$\begin{bmatrix} 4 & -1 & -3 \\ 0 & 9 & -3 \\ 0 & 0 & 2 \end{bmatrix} \equiv \mathbf{R}$$

Direct computation will now confirm $\mathbf{R}\,\mathbf{R}^H = \mathbf{A}$.